Applying Participatory Design Theory to Designing Evaluation Methods

Diana Kusunoki
The iSchool at Drexel University
3141 Chestnut St.,
Philadelphia, PA 19104-2875 USA
diana.s.kusunoki@drexel.edu

Aleksandra Sarcevic
The iSchool at Drexel University
3141 Chestnut St.,
Philadelphia, PA 19104-2875 USA
aleksarc@drexel.edu

CHI '12, May 5–10, 2012, Austin, TX, USA.
ACM 978-1-4503-1016-1/12/05.

Abstract
System evaluators face several challenges in designing evaluation methods, including measurement and relevance, context, establishing common ground with users, and eliciting users' tacit knowledge. To address these challenges, we propose applying participatory design theory to designing evaluation methods by increasing user involvement and by integrating this process into the overall process of system design.

Keywords
Evaluation; Evaluation design; Participatory design.

ACM Classification Keywords
H.5.2 [Information Interfaces And Presentation]: User Interfaces - Evaluation/Methodology;

General Terms
Design; Human factors.

Introduction
There is a critical need for increased user involvement in the design of evaluation methods. The design paradigm in the field of HCI has shifted several times, but evaluation is still grounded in traditional theories. Evaluators now face multiple issues that make designing evaluation methods challenging. In this work,

we hypothesize that applying participatory design theory to evaluation design can reduce the impact of these issues on evaluation outcomes. We discuss these issues and present our plans for designing an evaluation method using participatory design and applying it in the advanced trauma care domain as part of the design process of information displays to support time-critical teamwork. In the following sections we review the paradigm shifts and evaluation practices in HCI to demonstrate (1) the lack of focus on evaluation, (2) the dependence on traditional evaluation methods, and (3) the feasibility of adapting participatory design theory to the development of evaluation methods.

Paradigm Shifts in HCI

The field of HCI has experienced several paradigm shifts over the past 30 years. Bannon [1] first described an initial paradigm shift "from human factors to human actors." Up until that point, the field had changed from a focus on training humans to operate machines, to designing machines to fit users, to understanding users' cognitive processes, and then to larger units of analysis from the individual to groups and organizations. Bødker [2] identified another shift in design practices that presented new challenges for second paradigm HCI practices. Important characteristics of this shift included the increase in the complexity, multiplicity, and ubiquity of technologies; the need to account for context; and, the attempt to understand the user experience, emotions, and reflexivity. Maceli and Atwood [8] extended Bannon and Bødker's discussions by suggesting that there is an emerging paradigm shift to meta-design where users are the designers, or "human crafters."

It appears that the paradigm shifts in HCI have mainly been described in terms of design. Evaluation is often discussed as a sort of necessary evil—an assumed, yet taken for granted process. The design and description of evaluation methods are frequently left to readers' imagination. Because our designs and their validation in the community depend heavily on evaluation [6], it appears that evaluation merits its own design process and needs to be treated as important as the design process itself. However, design and evaluation are often separated in practice [3,11]. It is therefore essential that evaluation be planned, implemented, and reported with rigor, and at the same time, carefully woven into the system design process.

Practices and Perceptions of Evaluation in HCI

There have been many concerns raised about approaching evaluation in HCI. Students of HCI are taught that evaluation is important and should always accompany design. However, evaluation is typically a vaguely defined process. It still draws on traditional methods such as usability testing, heuristic evaluation, cognitive walkthrough, and task analysis. In many cases, evaluation methods from the first and second paradigms are still the standard for evaluating designs conceived through third paradigm principles.

At the same time, systems are becoming increasingly dynamic and complex, requiring evaluation methods that can grow and adapt to the context in which systems are being used. Yet, it may not be necessary to dismiss older paradigm evaluation methods like usability testing or heuristic evaluation if they are found to be the appropriate tools [2,3]. If we do find that they are not suitable for evaluating our designs, it is

important that we pursue methods that fit both designs and users in their context [3].

A few researchers have noted the potential harm in relying on discount evaluation methods (i.e., heuristic evaluation, basic paper prototypes, and user testing with a few users) because they may give us false sense of rigor, meaningless or trivial results, and misdirect design directions [6]. It is also easy to believe that evaluation methods that are popular or that have been validated by other researchers can be effectively implemented anywhere, anytime regardless of whether they fit or not. This type of evaluation methodology has been likened to "drunks under the lamppost" searching only in the light for their lost keys [3].

Evaluation is more about *how* to carefully integrate selected methods into the system design process than just *what* methods to select. Regardless of the methods developed or used, overarching cohesion should exist concerning the methodologies supporting the design and evaluation of the system. We next discuss the challenges in designing effective evaluation methods.

Challenges in Designing Evaluation Methods

The design of evaluation methods has not always been a structured or intentional process. We need better ways to (1) methodically design evaluation methods, (2) cope with varying contexts, (3) reach mutual understanding, (4) communicate with and elicit feedback from users, and (5) integrate evaluation design into the system design process.

Measurement and relevance. Ensuring that our evaluation techniques and instruments are accurately capturing the most important aspects of the system is difficult. The main factor influencing the quality of the information elicited through interviews or surveys is question design. Questions mediate between the intentions of the researchers and the understanding and attitudes of users. Many of the design principles discussed in the survey methodology literature can be applied to the design of evaluation instruments. These can help shape our understanding for how to design, implement, and analyze surveys to gather the most useful, representative, and appropriate information.

Context. Accounting for users' various backgrounds, environments, and relationships can be challenging in evaluation. Deciding what to evaluate and how is usually left to the evaluator. Problems that arise are often "complex, ill-structured, or wicked," and we should thus situate both design and evaluation in the context of use [11]. There are important considerations for ensuring that design and evaluation methods align. Taking time to understand users, identify assumptions made, and involve users in the evaluation design process may minimize the effects of evaluators' inherent biases.

Common ground. Users (and often evaluators themselves) can have different, even conflicting ideas about what makes a design successful [7]. It is important to identify and negotiate differences in expectations, priorities, and use practices to reach mutual understanding. Moreover, without first grounding the language as the basis for discussion, communication among users, researchers, and other stakeholders is problematic. Aligning user and evaluator interpretations of terminology can help evaluation proceed more efficiently and reliably.

Eliciting tacit knowledge. Users are typically viewed as key informants in evaluation. However, it is not always easy for users to conceptualize and articulate their knowledge. Eliciting knowledge that is implicit or tacit in the minds of users is a great challenge [4]. Dealing with tacit knowledge introduces complex issues with regard to language, interpretation, and measurement. It is difficult to determine how tacit knowledge can be measured and if any meanings and interpretations made by users and evaluators actually match.

Integrating evaluation design into the system design process. As discussed earlier, the design process is often separated from evaluation in practice [3,11]. This separation can yield wasted evaluation efforts or produce lower quality data. Developing and conducting an evaluation process after the system is designed also risks missing the mark, requiring more time and effort to redesign the system. If the evaluation process itself is not iteratively refined before finalizing the design, evaluators also risk overlooking valuable feedback critical to system design and performance measures.

Moving Evaluation Forward

To address the core evaluation challenges discussed above, we propose taking a participatory design approach to the design of evaluation methods. Participatory design [9] is an appropriate methodology for designing evaluation methods because it not only promotes a high level of user involvement, but also sees the designer as an essential field guide [6]. While it is crucial that users be involved in the evaluation design process, it is also important to have evaluator expertise in developing targeted instruments. Participatory design and user-centered design are closely related, but are fundamentally different approaches to design. User-centered design focuses on designing for the user, but might not necessarily be participatory where users are actively involved in the design process [11].

By including users in the evaluation design process, we can learn what users value and confirm that users and evaluators understand each other. This involves understanding users' context, establishing common ground, and collectively determining the understandability, appropriateness, and content of questions. We recommend integrating this participatory evaluation design process into the overarching participatory system design process.

Feasibility Assessment Plans: Application in the Trauma Resuscitation Domain

To assess the feasibility and effectiveness of our approach, we plan to apply this participatory evaluation design process in the advanced trauma care domain as part of the design process of information displays to support time-critical teamwork. Advanced trauma care is a highly complex, dynamic, and time- and safety-critical domain. The nature of the environment presents challenges for care providers' decision-making using complex information and information sources. Designing and evaluating systems to support providers' information needs is subsequently difficult. Evaluation design challenges apply to this domain, but are intensified due to the nature of the trauma environment and limitations on participant access and availability. System evaluation in this domain therefore needs to be flexible but targeted, efficient, and rigorous.

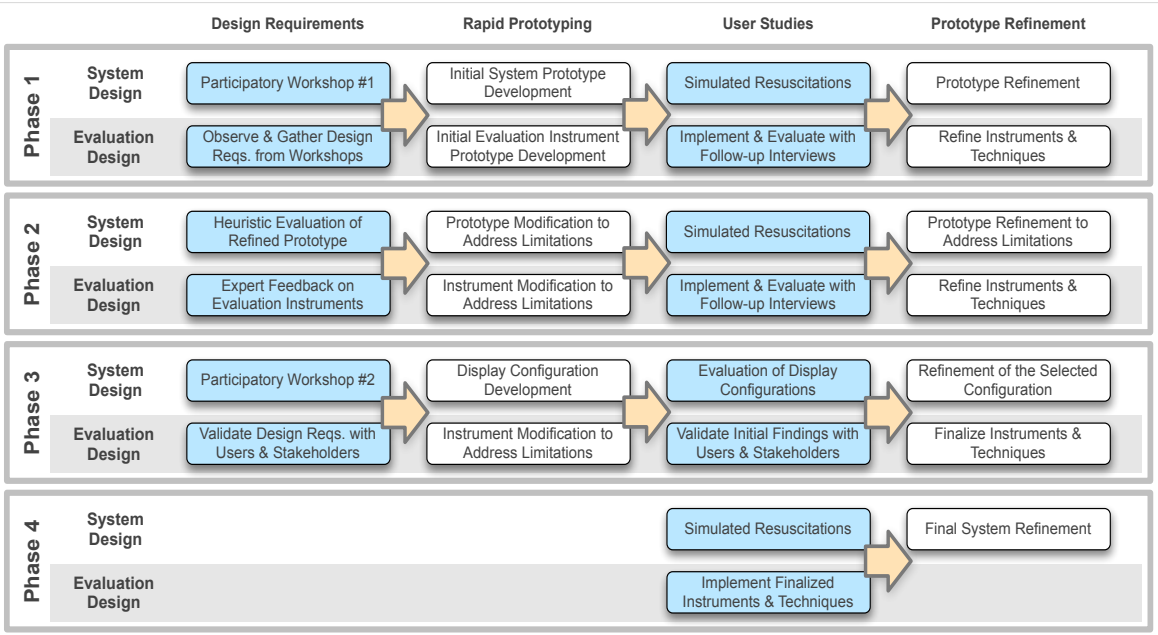

Figure 1: Integrating participatory evaluation design into the iterative system design process.

PARTICIPATORY EVALUATION DESIGN STEPS

We will integrate a structured evaluation design process into a participatory and iterative system design process of an information display to support providers' information needs during trauma resuscitation (Figure 1). Building on extensive observations and fieldwork, the display design process (top row of each phase in Figure 1) will involve providers from different disciplines and will focus on the types of information to display, how to display information, and the user experience. The goal of our research presented here is to integrate a participatory evaluation design process into this participatory design process for the information display (bottom row of each phase in Figure 1). This integration will be done as follows. While identifying system requirements with users, we will examine what users feel should be evaluated about the system for it to support their needs. When the system prototype is being designed, the evaluation instruments and techniques will also be designed based on the requirements gathered. The instruments will be evaluated based on user studies and interviews immediately following the system evaluation. Finally, when redesigning the system, the evaluation method will be refined based on user feedback and experiences during implementation. We next describe these evaluation design steps in greater detail.

Design requirements gathering and rapid prototyping: We will observe users during participatory workshops and user studies to distinguish between what we already know from what we want to know about users. This will reveal behaviors, factors, and barriers that we did not consider before, or that are contrary to our previous assumptions and findings [10]. Observations will also help us to understand users' context, establish common ground, and identify important terminology for better communication. Data gathered through participatory workshops will also be used for creating an initial set of evaluation design requirements. We will then design instruments and protocols to better represent what users do. We believe this process will help us avoid asking questions that do not apply or that will not reveal new information about users.

User studies, prototype refinement, and validation: We plan to conduct semi-structured interviews during user studies to gain insight into users' understanding and perceptions about the system. We will use concurrent

and retrospective interview techniques to verify and enrich the information provided by users [10]. Asking users basic follow-up questions about their answers will allow them to clarify vague responses and elaborate on answers that may lead to new insights [5]. During later participatory workshops (Phase 3 in Figure 1), users will be asked to retrospectively confirm findings from previous workshops and interviews, reducing the need for separate follow-up sessions. While conducting heuristic evaluation of the system design, we will gather expert feedback on the technical development of the questions and instruments. This will help ensure that survey design recommendations are followed. Evaluation instruments and techniques will be iteratively refined after receiving feedback from each of the workshops, user studies, and heuristic evaluations.

Conclusion and Limitations

This paper makes the argument for participatory evaluation design and describes how this process can be integrated into the system design process. Working with evaluators and users from various disciplines with differing practices, foci, and standards for success can introduce difficulty into the evaluation design process. Participatory design does not necessarily guarantee user buy-in or mutual commitment to each other's needs. Taking extra time to understand each other and establish common ground through participatory workshops can help alleviate some of these issues. Implementing highly tailored approaches can also leave us with evaluation designs that are not generalizable to other contexts or domains. However, the evaluation design process itself is potentially transferrable. The proposed research will allow us to test new ways to develop and conduct evaluation design in a more structured and intentional fashion.

Acknowledgements

We thank Dr. Michael Khoo, Dr. Michael Muller, and Dr. Randall Burd for their insightful comments.

References

[1] Bannon, L.J. From human factors to human actors: The role of psychology and human-computer interaction studies in system design. In J.M. Greenbaum and M. Kyng (Eds.), *Design at Work: Cooperative Design of Computer Systems*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 25-44, 1991.

[2] Bødker, S. When second wave HCI meets third wave challenges. *Proc. NordiCHI 2006*, (2006), 1–8.

[3] Cockton, G. What worth measuring is. *Proc. Int'l Workshop on Meaningful Measures: Valid Useful User Experience Measurement (VUUM) 2008*, (2008), 60-66.

[4] Forsythe, D.E. "It's just a matter of common sense": Ethnography as invisible work. *Computer Supported Cooperative Work* 8, 1-2 (1999), 127-145.

[5] Fowler, F.J. *Improving Survey Questions: Design and Evaluation*. SAGE, 1995.

[6] Greenberg, S. and Buxton, B. Usability evaluation considered harmful (some of the time). *Proc. CHI 2008*, ACM Press (2008), 111-120.

[7] Kaye, J. and Sengers, P. The evolution of evaluation. *alt.chi 2007*, (2007).

[8] Maceli, M. and Atwood, M. From human factors to human actors to human crafters. *Proc. iConference 2011*, (2011), 98–105.

[9] Muller, M.J. and Kuhn, S. Participatory design. *Communications of the ACM* 36, 6 (1993), 24–28.

[10] Presser, S. *Methods for testing and evaluating survey questionnaires*. John Wiley & Sons, 2004.

[11] Wania, C.E., Atwood, M.E., and McCain, K.W. How do design and evaluation interrelate in HCI research? *Proc. DIS 2006*, ACM Press (2006), 90-98.